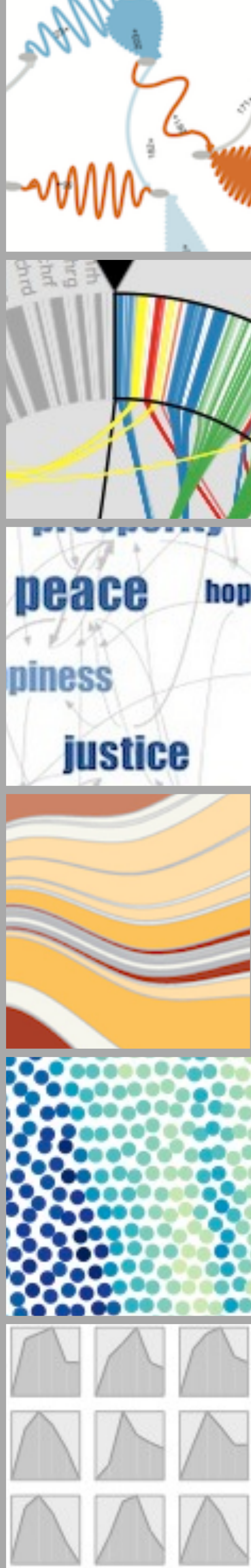
cs6964 | January 19 2012

DATA

Miriah Meyer

University of Utah

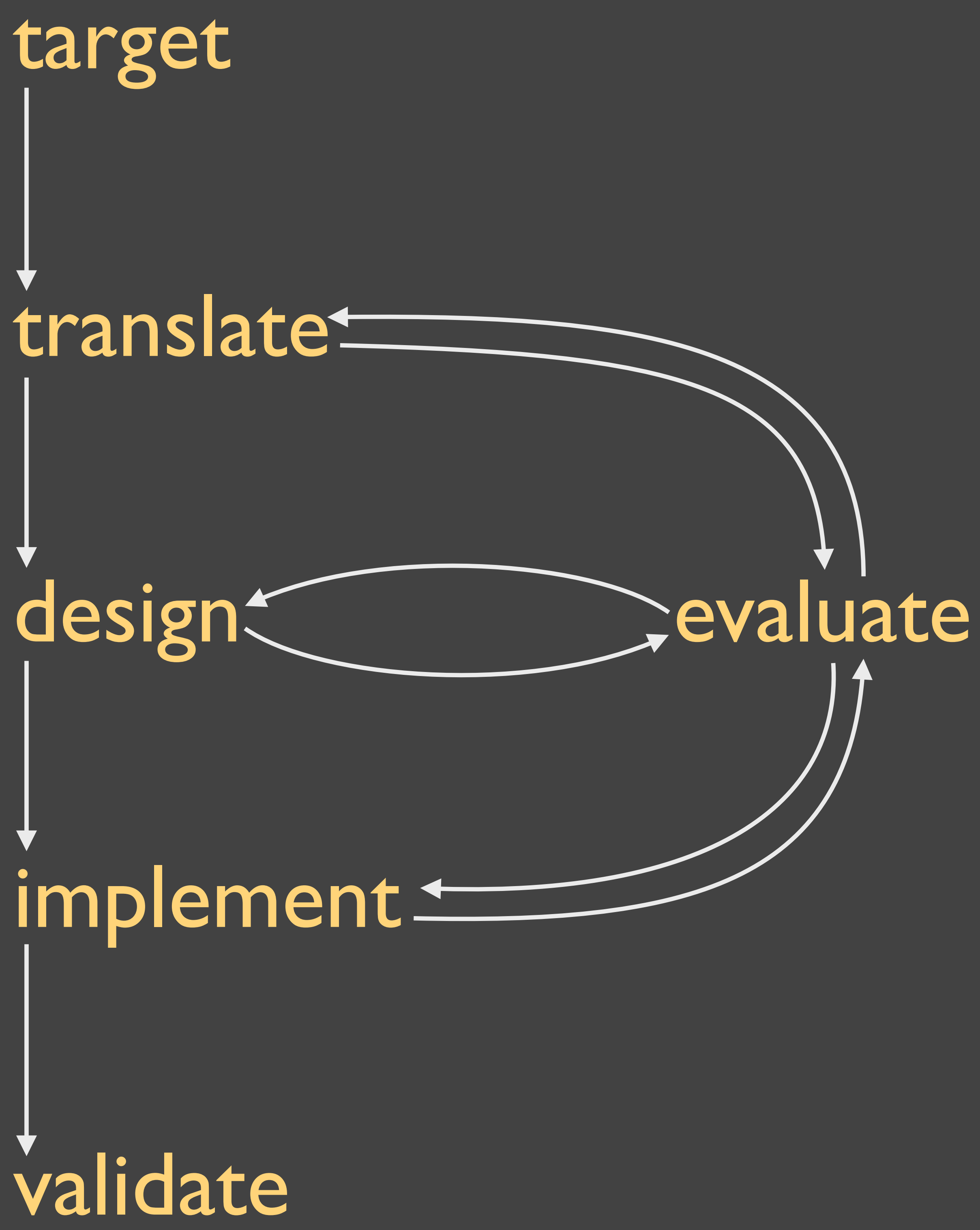


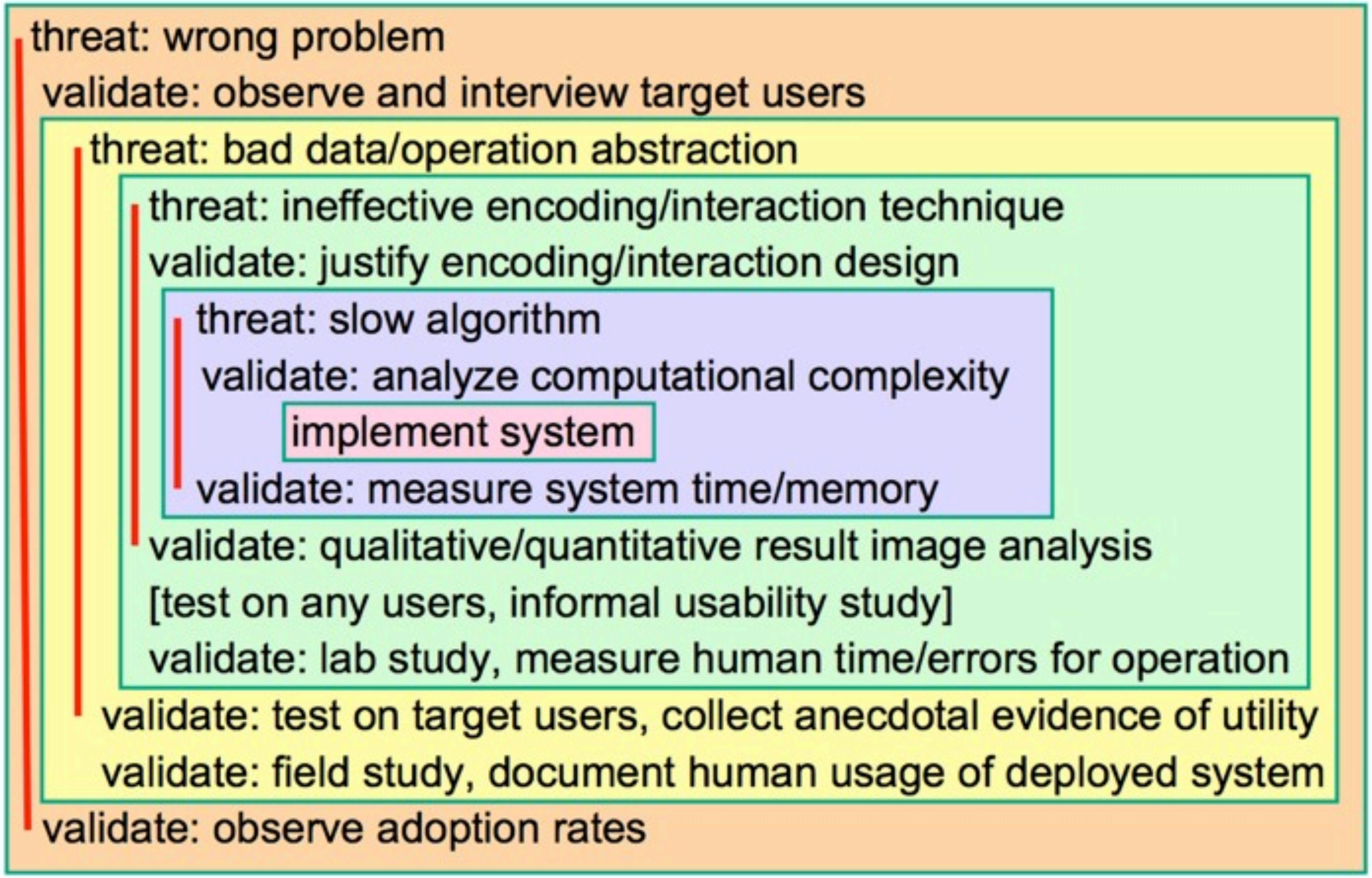
slide acknowledgements:

Tamara Munzner, University of British Columbia

Hanspeter Pfister, Harvard University

LAST TIME

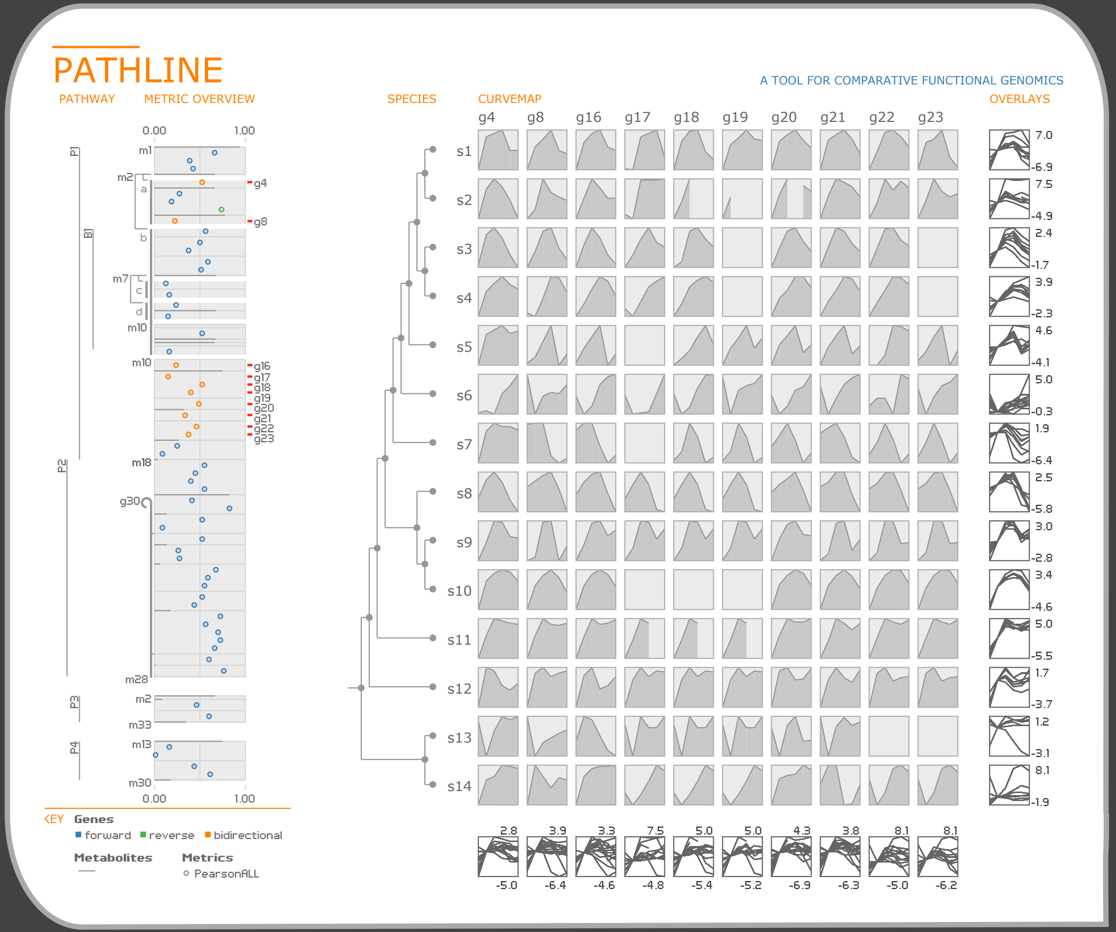

target
translate
design
evaluate
implement
validate



avoid validation mismatch

-cannot validate encoding with system timings
-cannot validate abstraction with lab studies

Munzner 2009



Pathline

A Tool for Comparative Functional Genomics Data

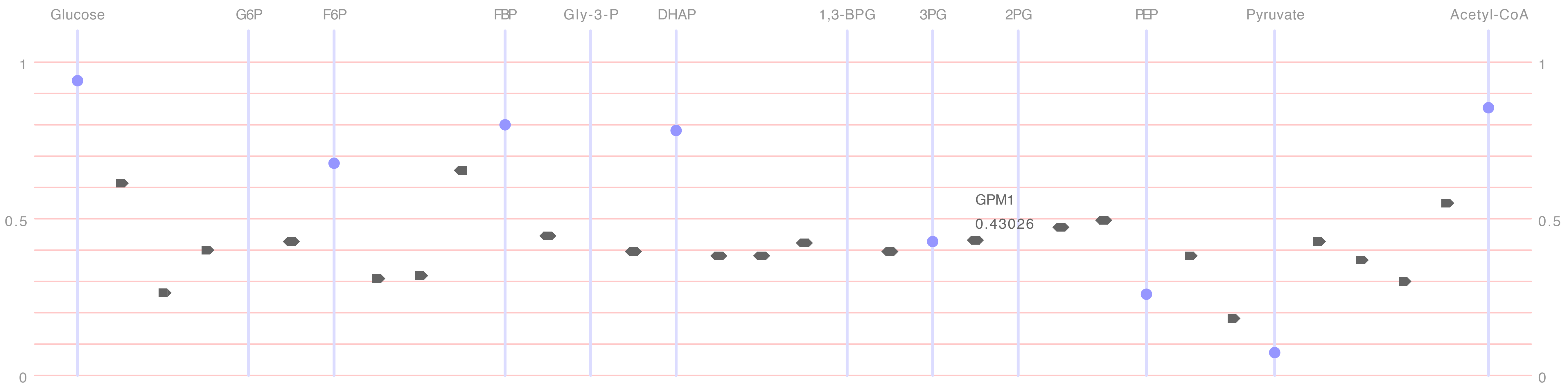
joint work with:

Bang Wong, Mark Styczynski, Tamara Munzner, Hanspeter Pfister

Pathline: A Tool for Comparative Functional Genomics
M. Meyer et al., IEEE/Eurographics EuroVis 2010.

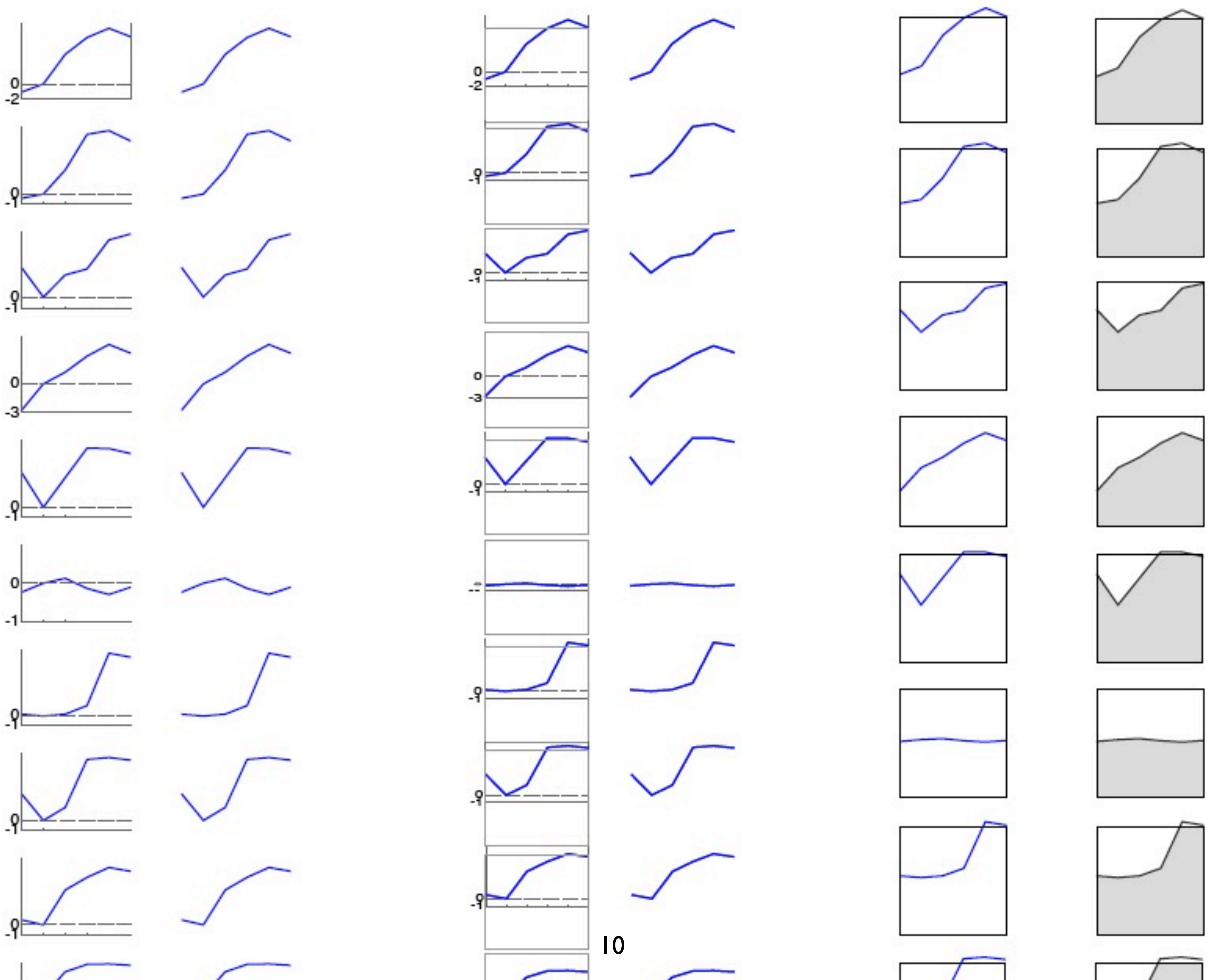
LESSONS LEARNED

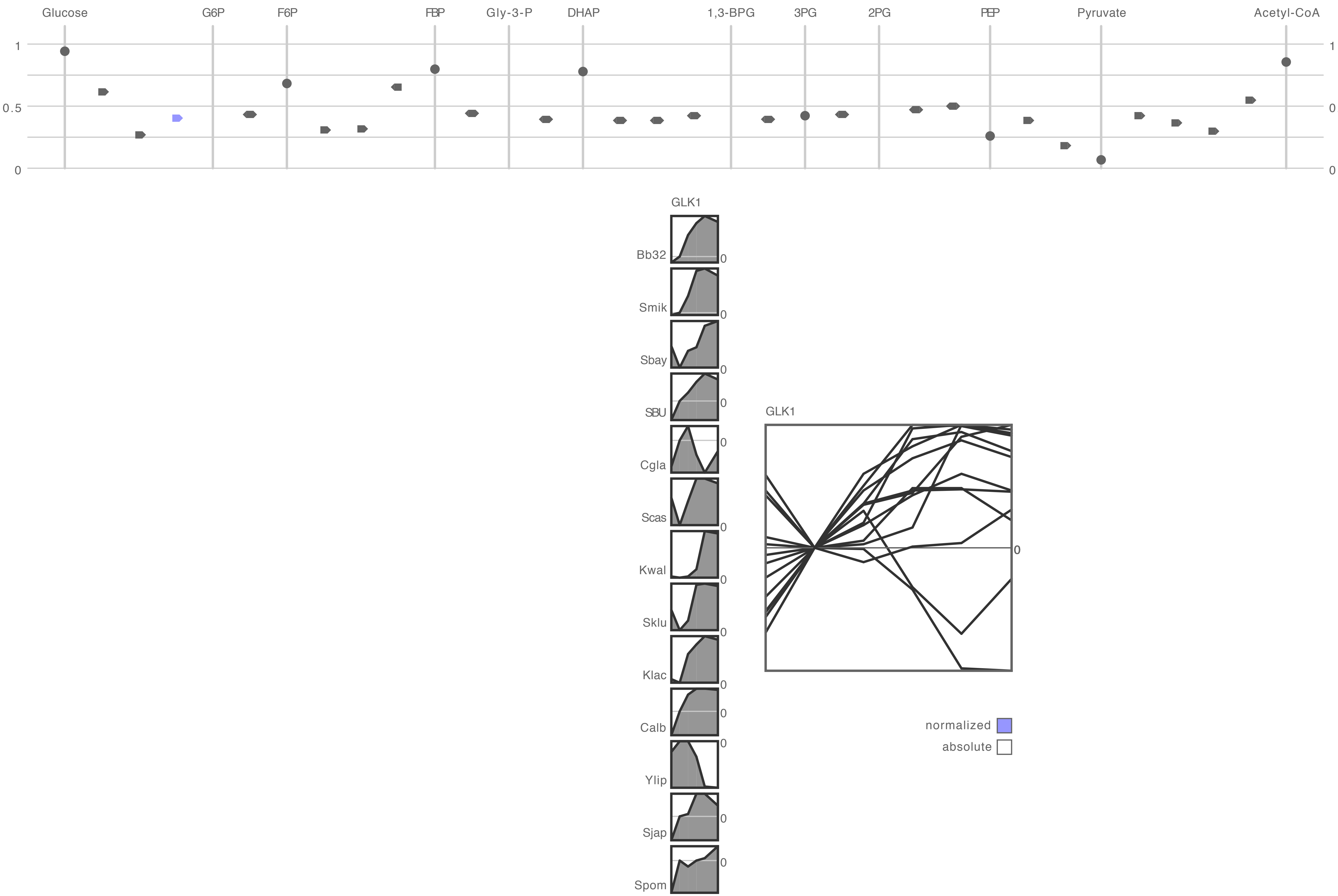
- **process supports efficient development**
- **collaborators' time commitment is front loaded**
- **rapid prototyping is essential**

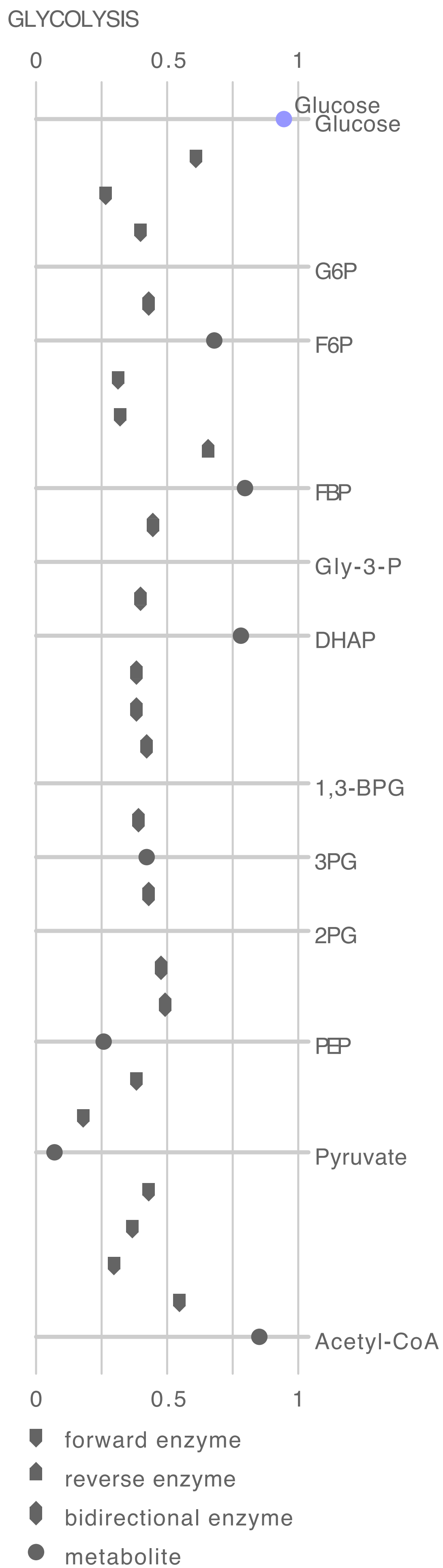

Glucose
G6P
F6P
FBP
Gly-3-P
DHAP
1,3-BPG
3PG
2PG
PEP
Pyruvate
Acetyl-CoA
1
0.5
0
GPM1
0.43026


Glucose
G6P
F6P
FBP
Gly-3-P
DHAP
1,3-BPG
3PG
2PG
PEP
Pyruvate
Acetyl-CoA
1
0.5
0
ENO2
0.49543

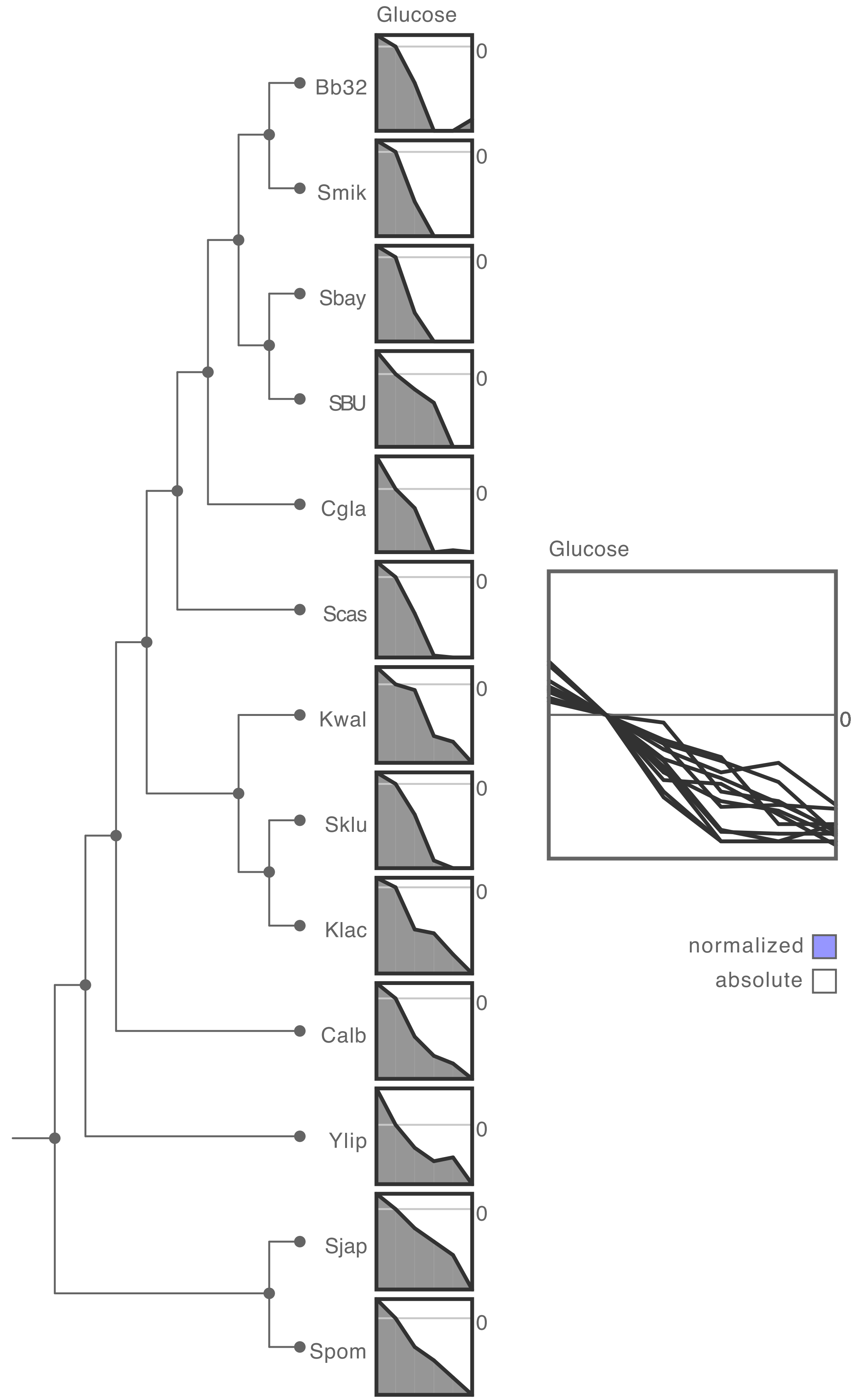
Relative scale | Absolute scale | Absolute scale, highlight pattern



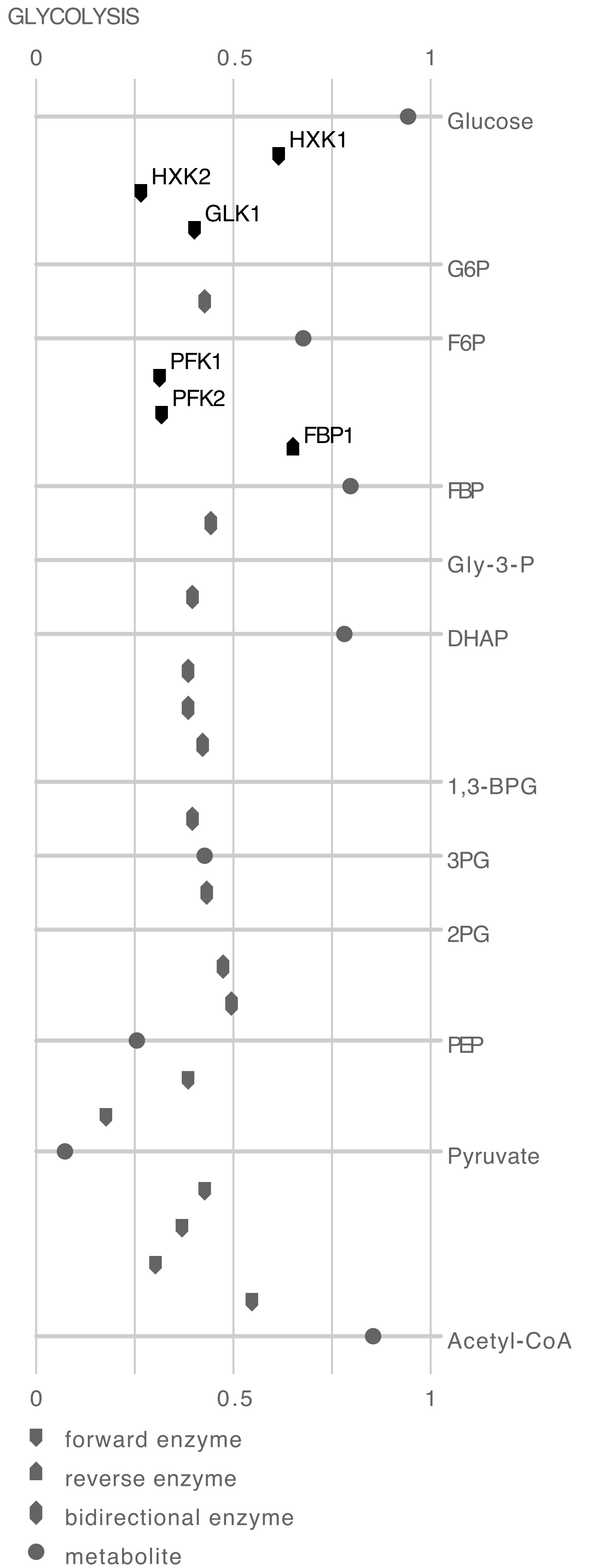

Glucose
G6P
F6P
FBP
Gly-3-P
DHAP
1,3-BPG
3PG
2PG
PEP
Pyruvate
Acetyl-CoA
1
0.5
0
GLK1
Bb32
Smik
Sbay
SBU
Cgla
Scas
Kwal
Sklu
Klac
Calb
Ylip
Sjap
Spom
GLK1
normalized
absolute

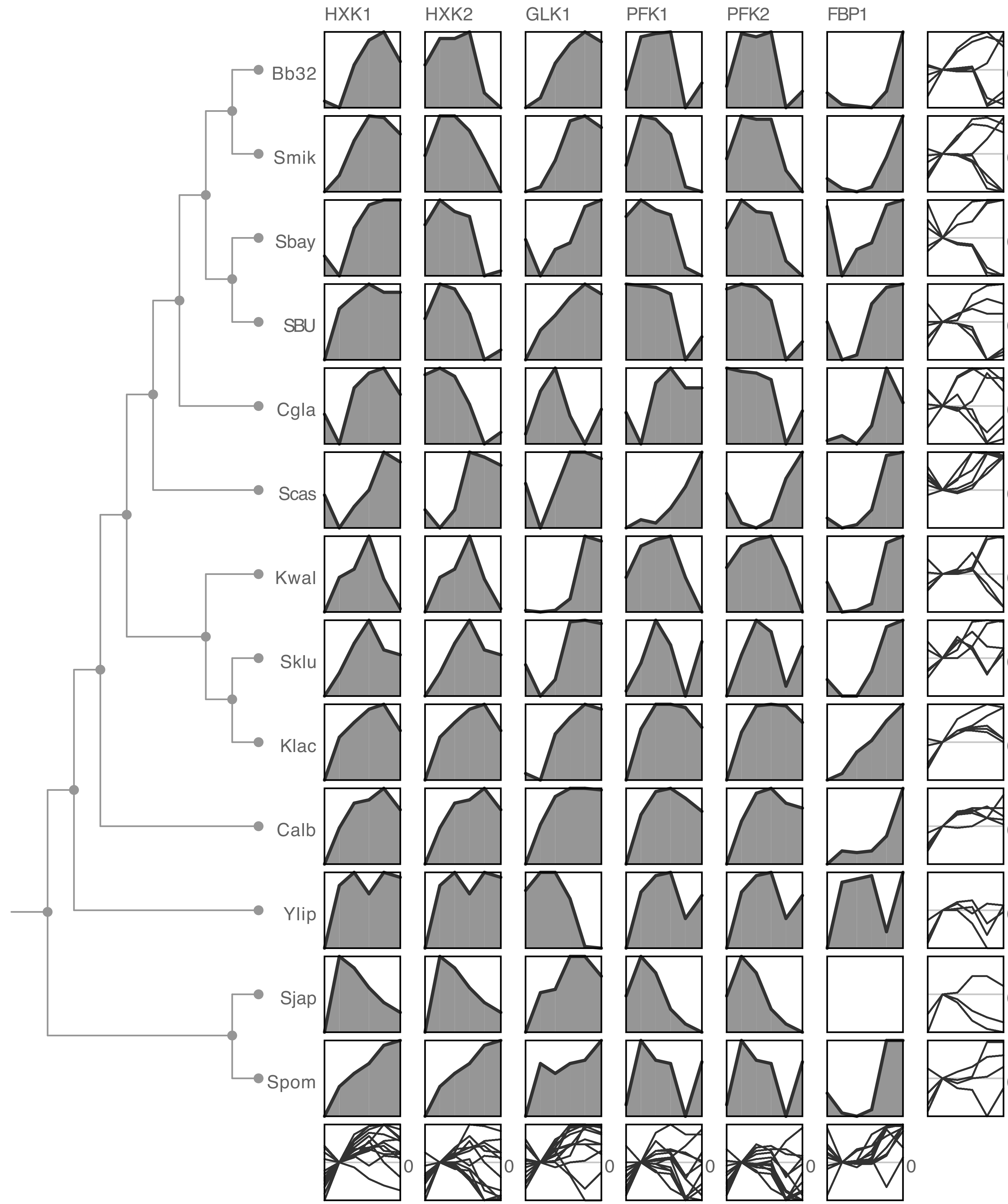


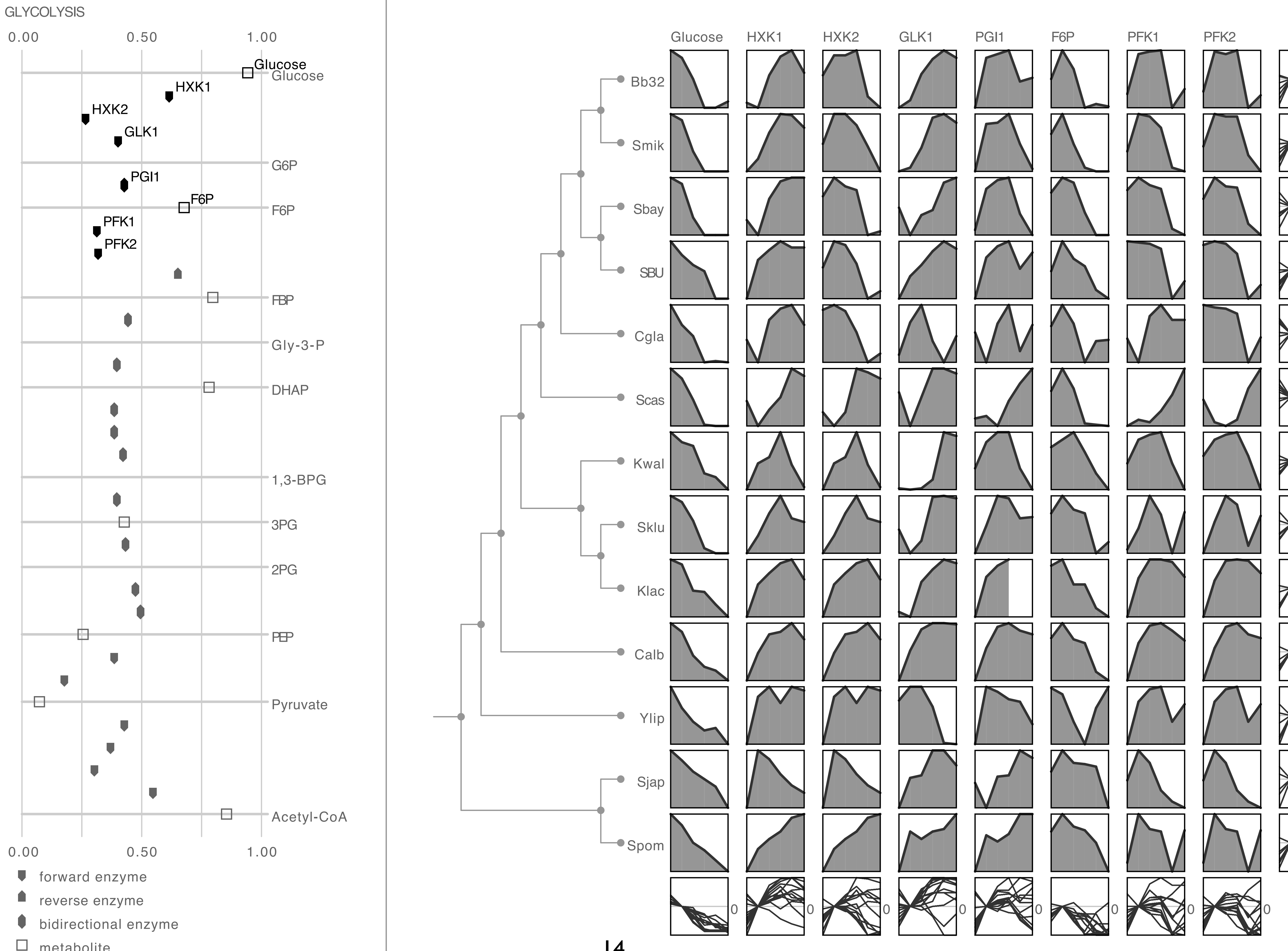
GLYCOLYSIS
0
0.5
1
Glucose
Glucose
G6P
F6P
FBP
Gly-3-P
DHAP
1,3-BPG
3PG
2PG
PEP
Pyruvate
Acetyl-CoA
forward enzyme
reverse enzyme
bidirectional enzyme
metabolite



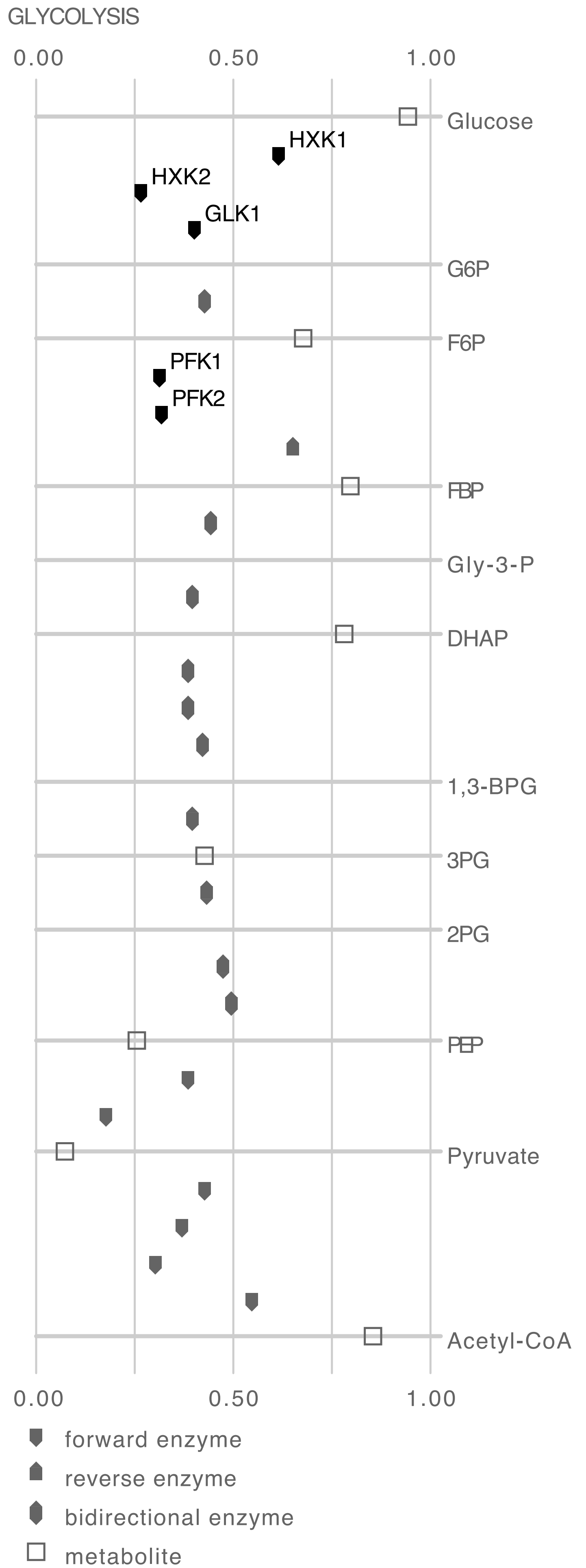
Glucose
Bb32
Smik
Sbay
SBU
Cgla
Scas
Kwal
Sklu
Klac
Calb
Ylip
Sjap
Spom
0
Glucose
normalized
absolute

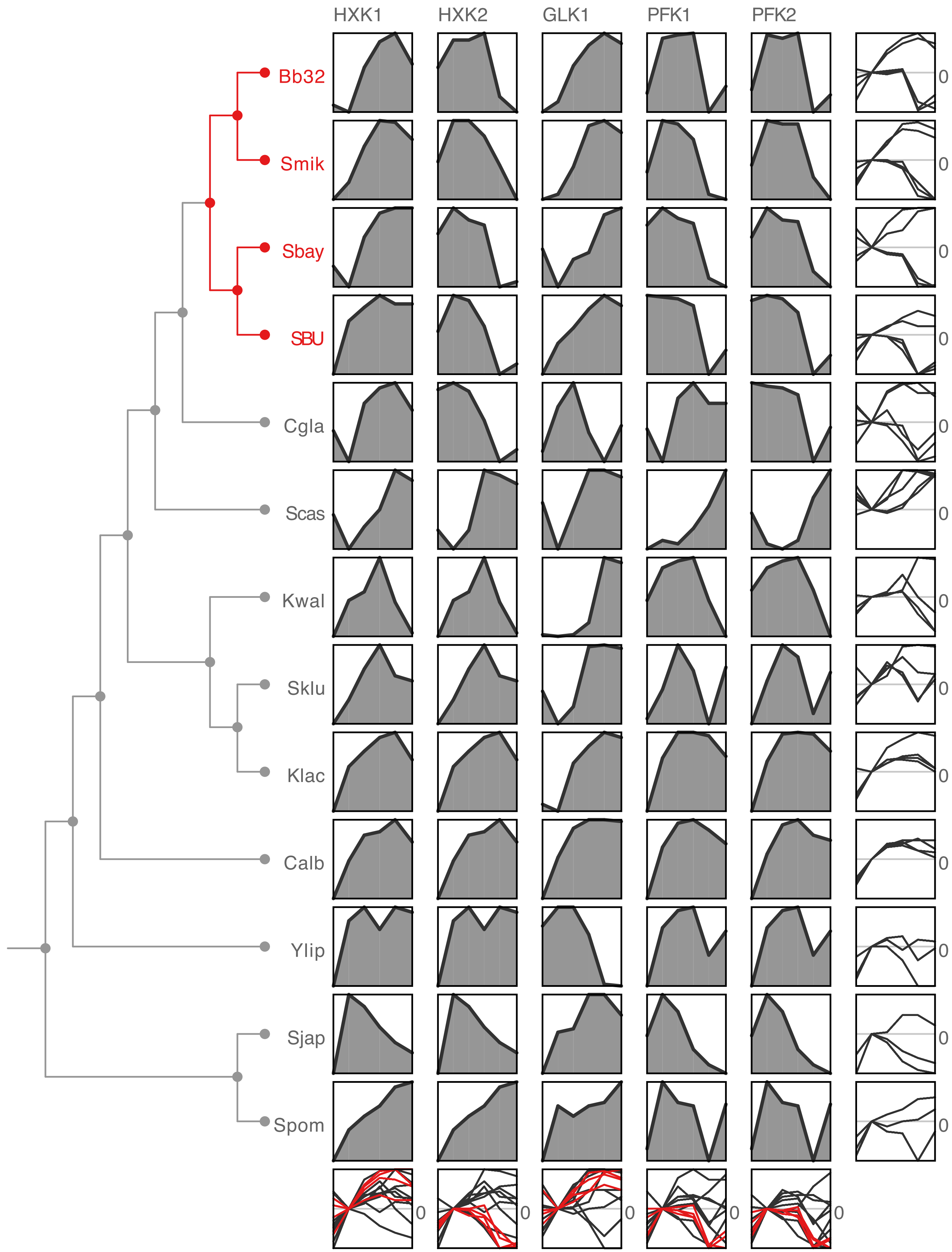

GLYCOLYSIS
0
0.5
1
Glucose
HXK1
HXK2
GLK1
G6P
F6P
PFK1
PFK2
FBP1
FBP
Gly-3-P
DHAP
1,3-BPG
3PG
2PG
PEP
Pyruvate
Acetyl-CoA
forward enzyme
reverse enzyme
bidirectional enzyme
metabolite

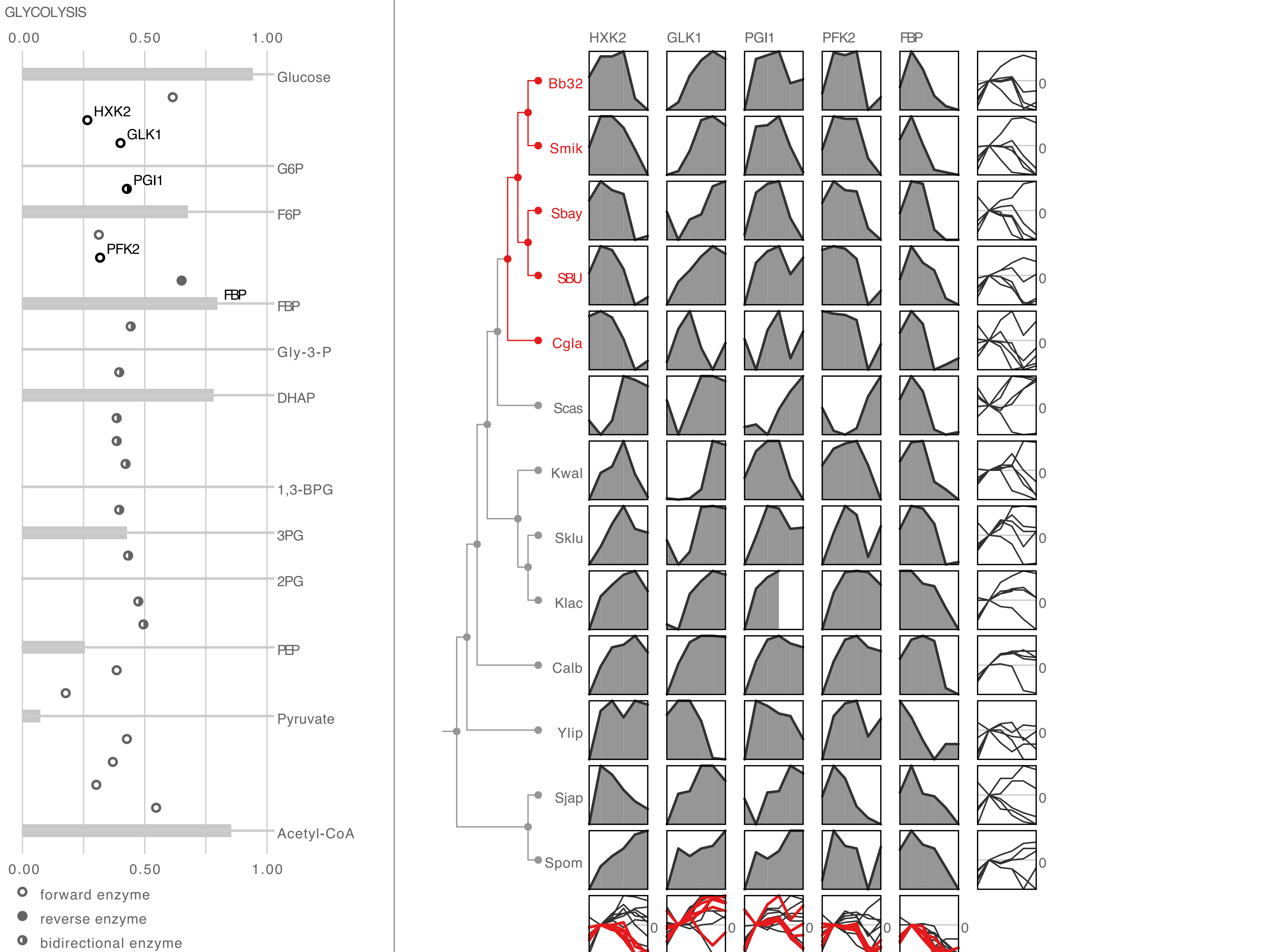

HXK1
HXK2
GLK1
PFK1
PFK2
FBP1
Bb32
Smik
Sbay
SBU
Cgla
Scas
Kwal
Sklu
Klac
Calb
Ylip
Sjap
Spom
0



GLYCOLYSIS
0.00
0.50
1.00
Glucose
HXK1
HXK2
GLK1
G6P
PGI1
F6P
PFK1
PFK2
FBP
Gly-3-P
DHAP
1,3-BPG
3PG
2PG
PEP
Pyruvate
Acetyl-CoA
forward enzyme
reverse enzyme
bidirectional enzyme
metabolite
Glucose
HXK1
HXK2
GLK1
PGI1
F6P
PFK1
PFK2
Bb32
Smik
Sbay
SBU
Cgla
Scas
Kwal
Sklu
Klac
Calb
Ylip
Sjap
Spom
0


GLYCOLYSIS
0.00
0.50
1.00
Glucose
HXK1
HXK2
GLK1
G6P
F6P
PFK1
PFK2
FBP
Gly-3-P
DHAP
1,3-BPG
3PG
2PG
PEP
Pyruvate
Acetyl-CoA
forward enzyme
reverse enzyme
bidirectional enzyme
metabolite


HXK1
HXK2
GLK1
PFK1
PFK2
Bb32
Smik
Sbay
SBU
Cgla
Scas
Kwal
Sklu
Klac
Calb
Ylip
Sjap
Spom
0


GLYCOLYSIS
0.00
0.50
1.00
Glucose
HXK2
GLK1
G6P
PGI1
F6P
PFK2
FBP
FBP
Gly-3-P
DHAP
1,3-BPG
3PG
2PG
PEP
Pyruvate
Acetyl-CoA
forward enzyme
reverse enzyme
bidirectional enzyme
HXK2
GLK1
PGI1
PFK2
FBP
Bb32
Smik
Sbay
SBU
Cgla
Scas
Kwal
Sklu
Klac
Calb
Ylip
Sjap
Spom
0

LESSONS LEARNED

- **process supports efficient development**
- **collaborators' time commitment is front loaded**
- **rapid prototyping is essential**
- **put off coding as long as possible**

contributions

- **Pathline**
  - multiple genes, time points, species, and pathways
- **linearized pathway representation**
- **curvemap**
- **tool deployment**
  - open source
  - used daily by several collaborators

- visualization design process
- **types of research contributions**

5.- PAPER TYPES

A Visweek paper typically falls into one of five categories: technique, system, design study, evaluation, or model. We briefly discuss these categories below. Although your main paper type has to be specified during the paper submission process, papers can include elements of more than one of these categories. Please see "Process and Pitfalls in Writing Information Visualization Research Papers" by Tamara Munzner for more detailed discussion on how to write a successful Visweek paper.

Technique papers introduce novel techniques or algorithms that have not previously appeared in the literature, or that significantly extend known techniques or algorithms, for example by scaling to datasets of much larger size than before or by generalizing a technique to a larger class of uses. The technique or algorithm description provided in the paper should be complete enough that a competent graduate student in visualization could implement the work, and the authors should create a prototype implementation of the methods. Relevant previous work must be referenced, and the advantage of the new methods over it should be clearly demonstrated. There should be a discussion of the tasks and datasets for which this new method is appropriate, and its limitations. Evaluation through informal or formal user studies, or other methods, will often serve to strengthen the paper, but are not mandatory.

System papers present a blend of algorithms, technical requirements, user requirements, and design that solves a major problem. The system that is described is both novel and important, and has been implemented. The rationale for significant design decisions is provided, and the system is compared to documented, best-of-breed systems already in use. The comparison includes specific discussion of how the described system differs from and is, in some significant respects, superior to those systems. For example, the described system may offer substantial advancements in the performance or usability of visualization systems, or novel capabilities. Every effort should be made to eliminate external factors (such as advances in processor performance, memory sizes or operating system features) that would affect this comparison. For further suggestions, please review "How (and How Not) to Write a Good Systems Paper" by Roy Levin and David Redell, and "Empirical Methods in CS and AI" by Toby

DATA ABSTRACTION

comments on readings?

target

↓

translate

↓

design

↓

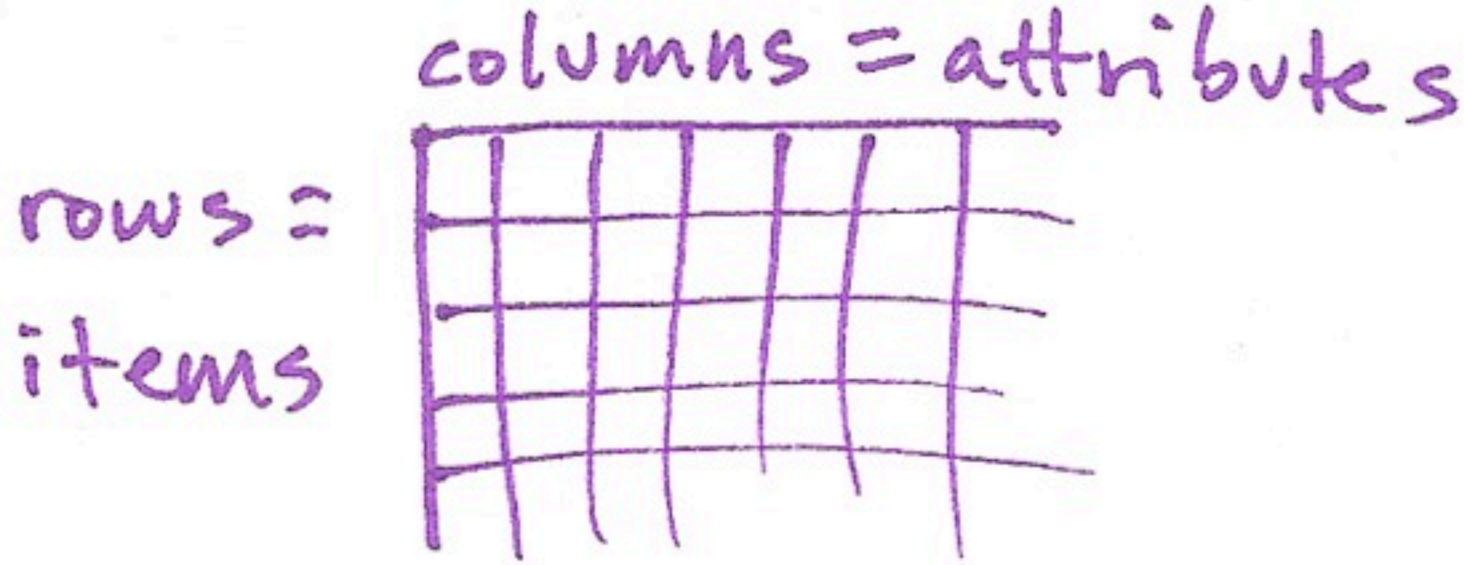
implement

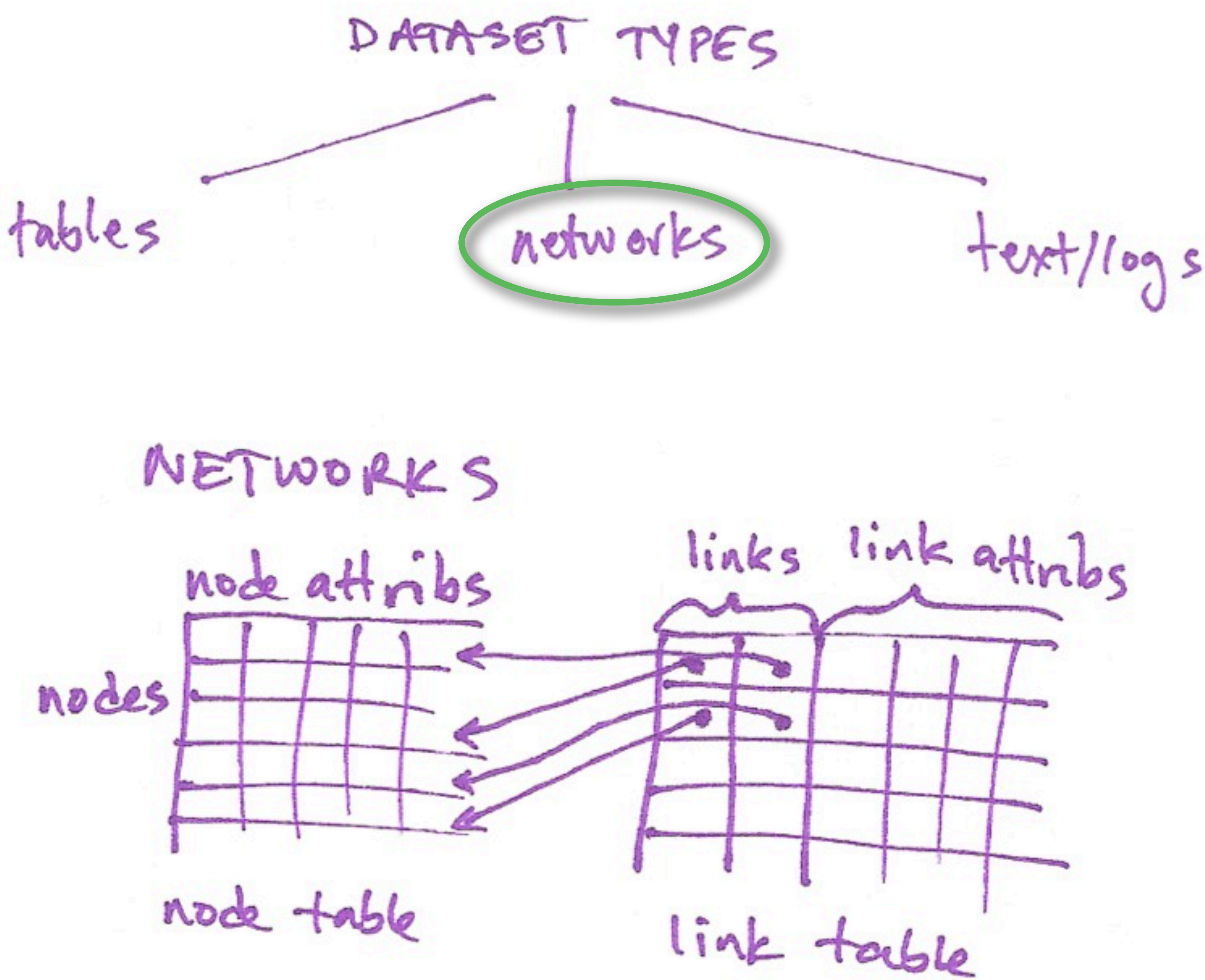
↓

validate

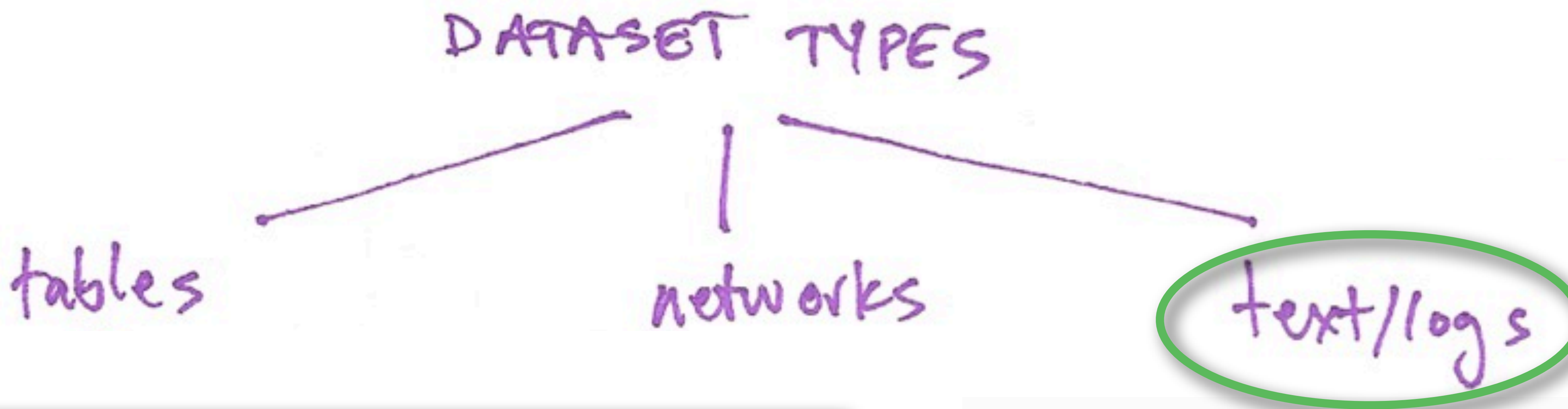
DATASET TYPES

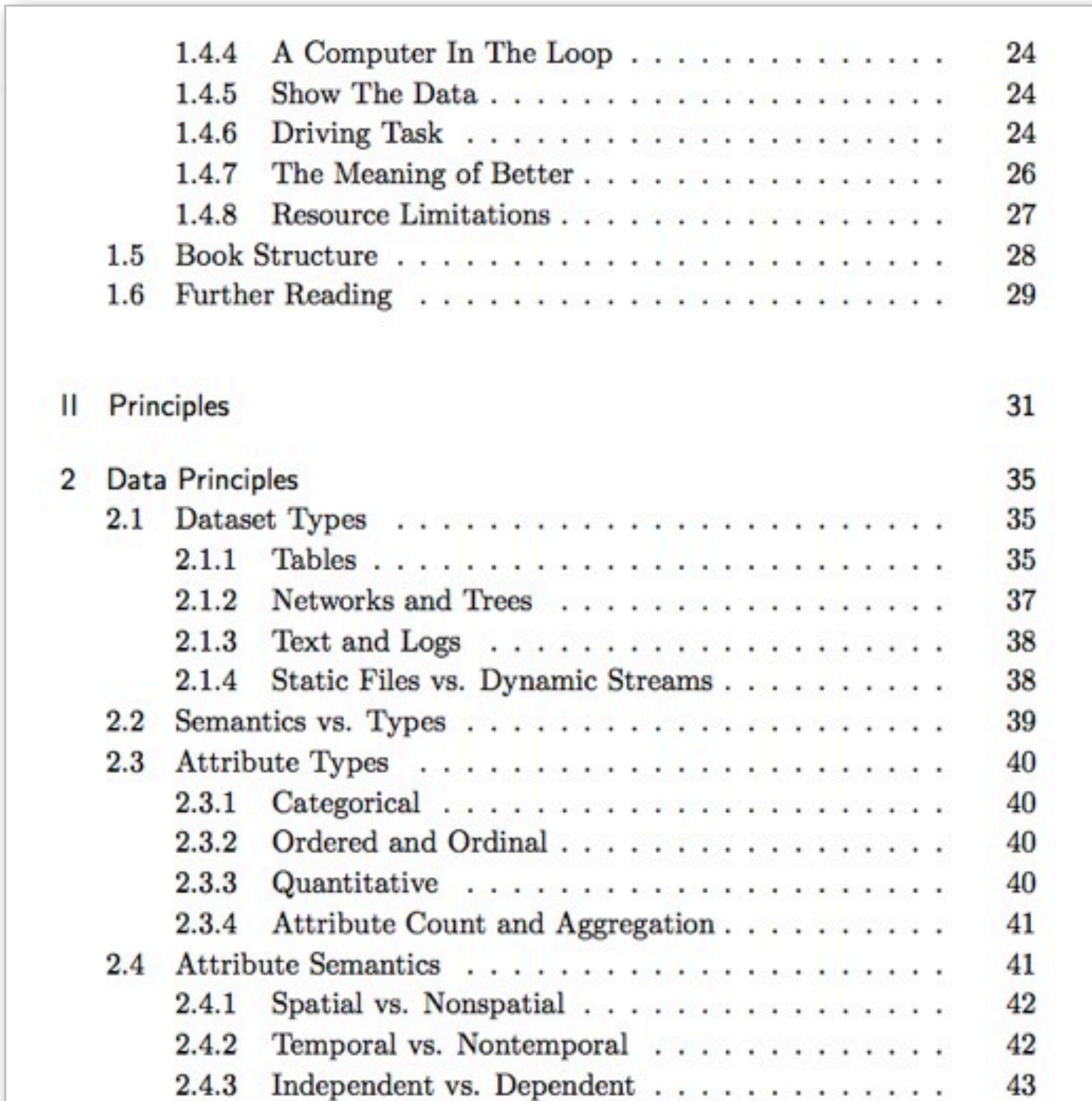
TABLES

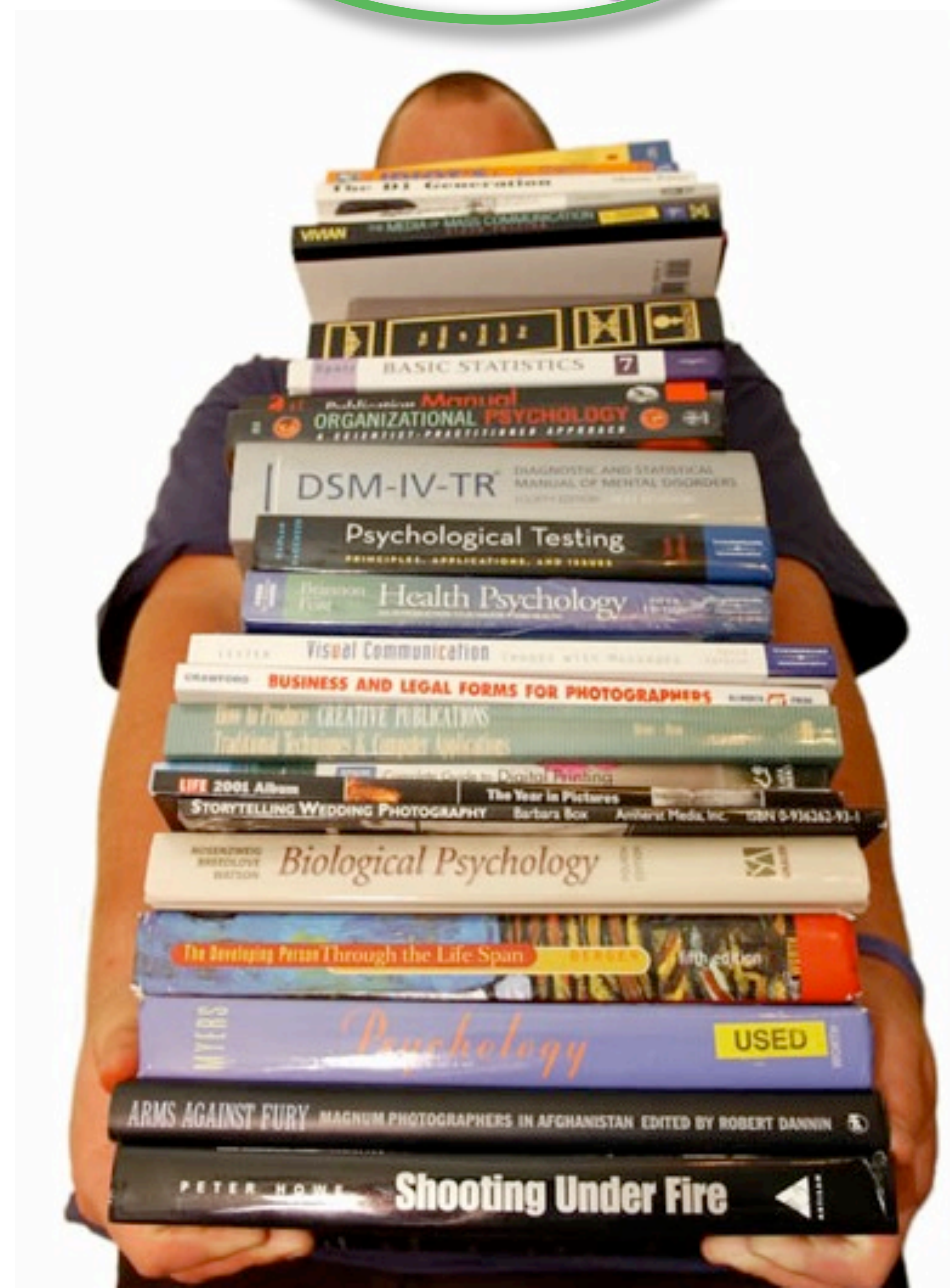


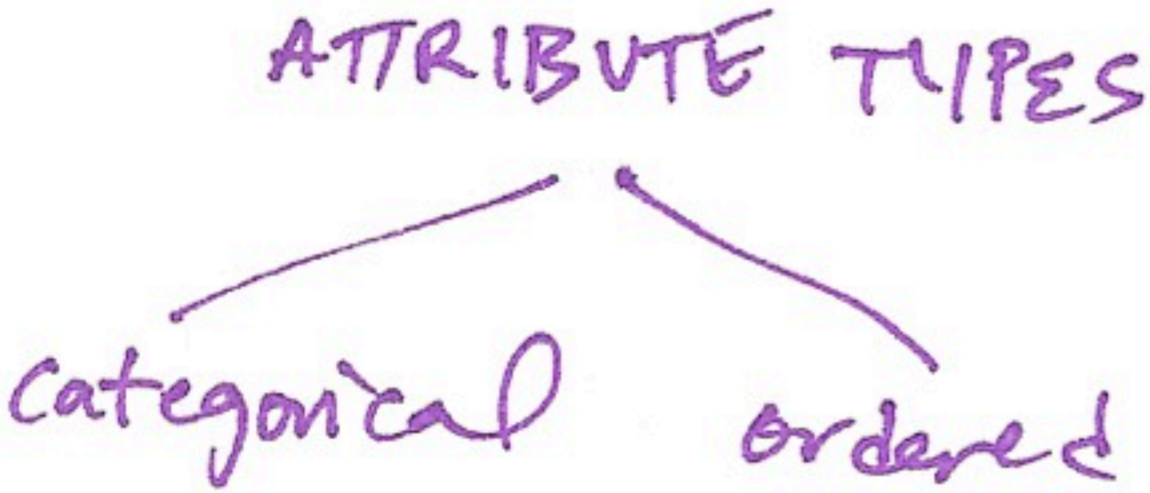


graph: another word for networks
tree: network without cycles


DATASET TYPES
tables
networks
text/logs




BASIC STATISTICS
ORGANIZATIONAL PSYCHOLOGY
DSM-IV-TR
Psychological Testing
Health Psychology
Visual Communication
BUSINESS AND LEGAL FORMS FOR PHOTOGRAPHERS
STORYTELLING WEDDING PHOTOGRAPHY
Biological Psychology
The Developing Person Through the Life Span
Psychology
USED
ARMS AGAINST FURY
PETER HOWE
Shooting Under Fire



mathematical interpretation

no implicit ordering

ATTRIBUTE TYPES

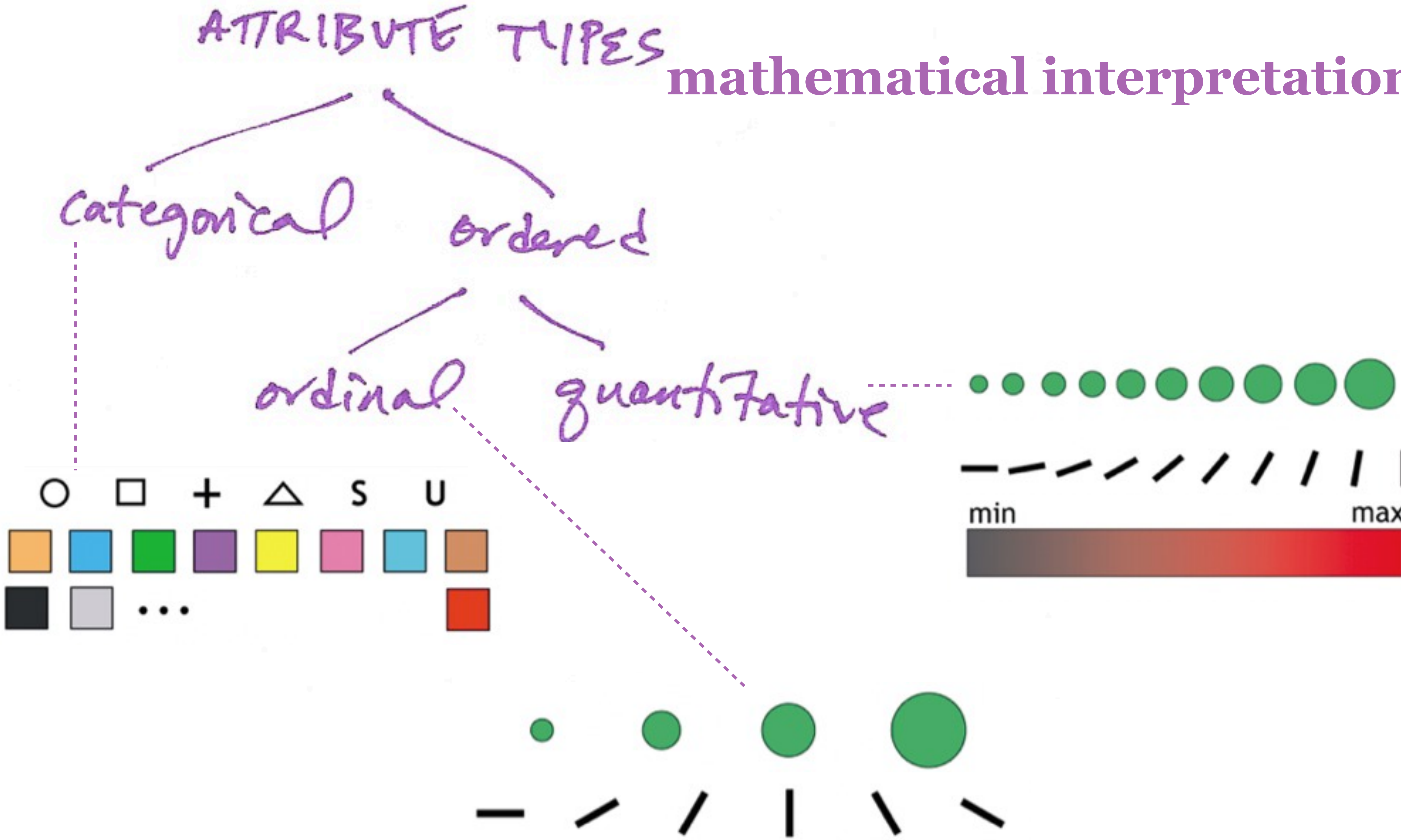
mathematical interpretation

- categorical
- ordered
  - ordinal
  - quantitative

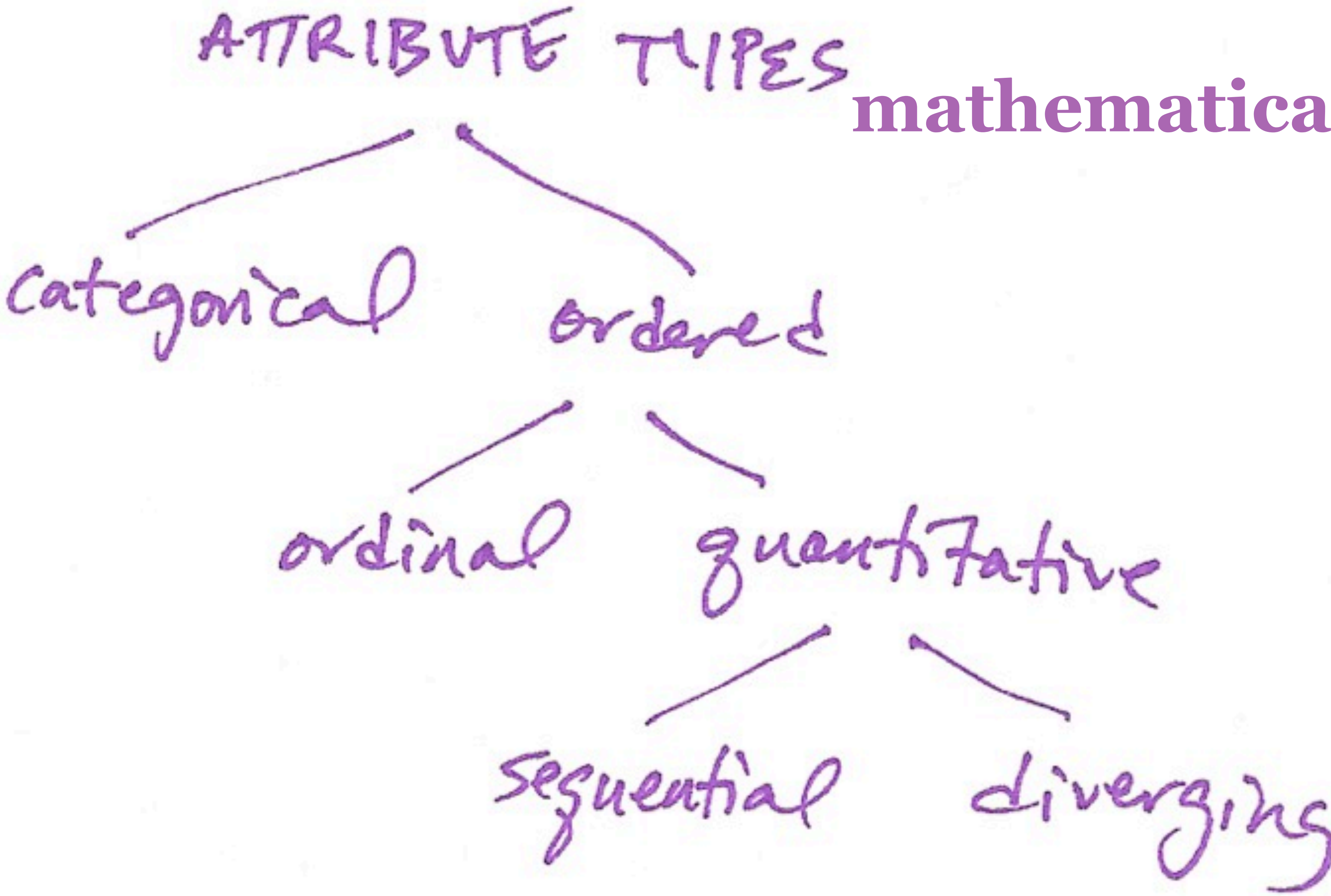
meaningful magnitude, can do arithmetic

examples?

mathematical interpretation



Stolte 2002



mathematical interpretation

examples?

A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box	0.72	7/17/07
32	7/16/07	2-High	Medium Box	0.6	7/18/07
32	7/16/07	2-High	Medium Box	0.65	7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
69	6/4/05	4-Not Specified	Small Pack	0.44	6/6/05
69	6/4/05	4-Not Spec		0.6	6/6/05
70	12/18/06	5-Low		0.59	12/23/06
70	12/18/06	5-Low		0.82	12/23/06
96	4/17/05	2-High		0.55	4/19/05
97	1/29/06	3-Medium		0.38	1/30/06
129	11/19/08	5-Low		0.37	11/28/08
130	5/8/08	2-High	Small Box	0.37	5/9/08
130	5/8/08	2-High	Medium Box	0.38	5/10/08
130	5/8/08	2-High	Small Box	0.6	5/11/08
132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
166	9/12/07	2-High	Small Box	0.55	9/14/07
193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08

quantitative
ordinal
categorical

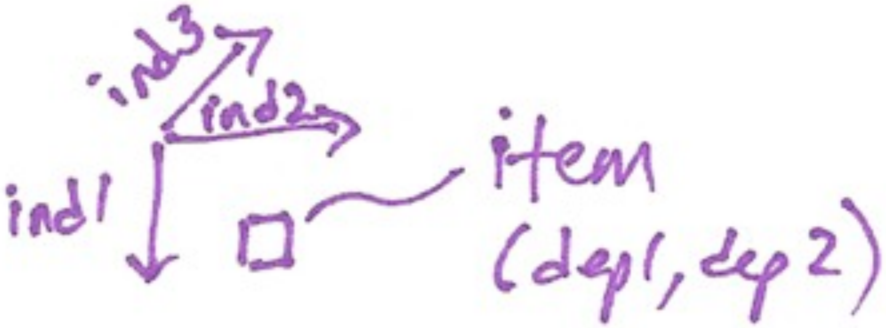
ATTRIBUTE SEMANTICS **real-world meaning**

ATTRIBUTE SEMANTICS **visualization specific**

- spatial/nonspatial
- temporal/nontemporal
- independent/dependent *
- continuous/discrete
- dimensions/measures

— independent/dependent

MULTIDIMENSIONAL TABLES



ind: index/independent attrib
dep: dependent/value attrib

SPATIAL FIELDS

depedent attrib count

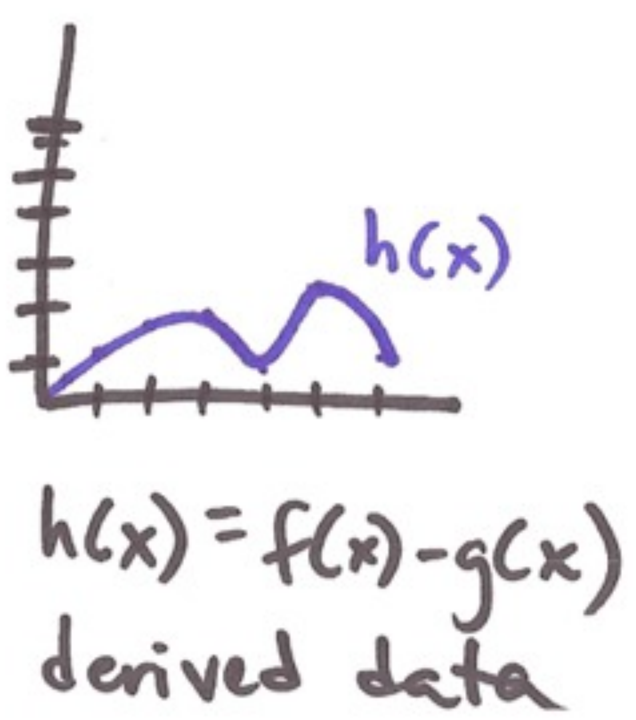
- scalar: 1 dep attrib
- vector: 2 dep attribs
- tensor: 3 dep

DATASET SEMANTICS **visualization specific**

- spatial/abstract
- static/timevarying

DERIVED ATTRIBUTES

- **derived attribute: compute from originals**
  - simple change of type
  - complex transformation
  - transformation is abstraction choice



DATA vs CONCEPTUAL MODEL

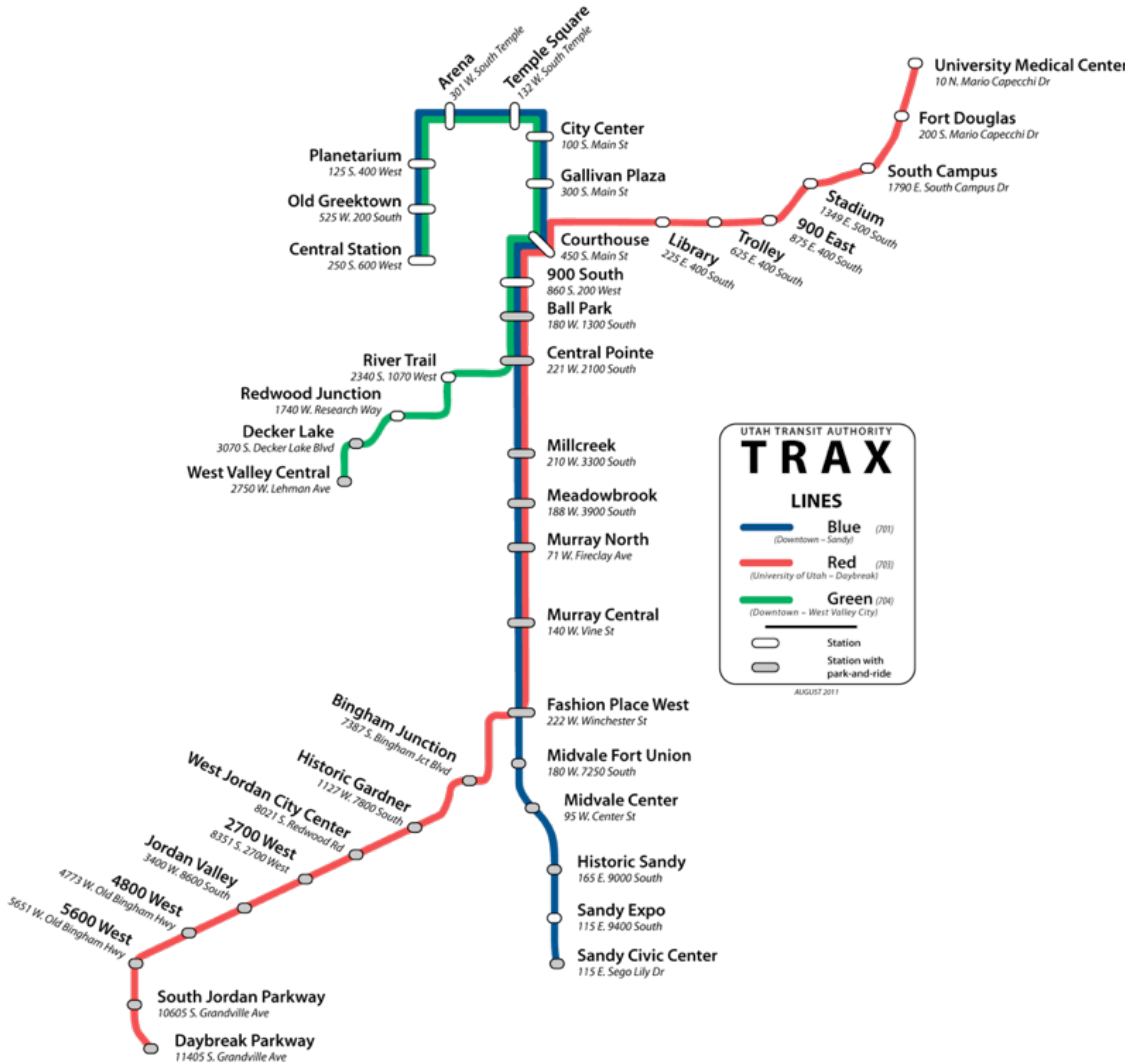
- **data model:** mathematical abstraction
  - set with operations, eg. floats with * / - +
- **conceptual model:** mental construction
  - includes semantics, supports reasoning
- **conceptual model motivates derived data**

EXAMPLE

- **from data model . . .**
  - 32.52, 54.06, -17.35, . . . (floats)
- **using conceptual model . . .**
  - temperature
- **to data type.**
  - continuous to 2 significant figures (Q)
  - hot, warm , cold (O)
  - above freezing, below freezing (C)

abstraction exercise . . .


ESPRESSO
& HOT TEA
GUATEMALA CASI CIELO 1.50 1.70 1.90
floral, lemon & cocoa
GUATEMALA CASI CIELO de-caf 1.50 1.70 1.90
floral, lemon & cocoa
SUMATRA 1.50 1.70 1.90
spicy, herbal & earthy
ESPRESSO 1.45 1.70 1.85
AMERICANO 1.60 1.80 2.00
CAFÉ LATTE 2.15 2.75 3.20
CAPPUCCINO 2.15 2.75 3.20
CAFÉ MOCHA 2.65 3.05 3.55
ORGANIC BREAKFAST 1.70 1.90 2.10
ORGANIC LONG LIFE GREEN TEA 1.70 1.90 2.10
MONSOON CHAI 1.70 1.90 2.10
CHAI TEA LATTE 2.40 2.95 3.35
BLACK TEA LATTE 2.20 2.55 3.20
HOT CHOCOLATE 2.50 2.75 3.00
HOMEMADE SYRUP FLAVORS .50 each


Arena
301 W. South Temple
Temple Square
132 W. South Temple
University Medical Center
10 N. Mario Capecchi Dr
Fort Douglas
200 S. Mario Capecchi Dr
City Center
100 S. Main St
Planetarium
125 S. 400 West
Gallivan Plaza
300 S. Main St
South Campus
1790 E. South Campus Dr
Old Greektown
525 W. 200 South
Stadium
1349 E. 500 South
900 East
875 E. 400 South
Central Station
250 S. 600 West
Courthouse
450 S. Main St
Library
225 E. 400 South
Trolley
625 E. 400 South
900 South
860 S. 200 West
Ball Park
180 W. 1300 South
River Trail
2340 S. 1070 West
Central Pointe
221 W. 2100 South
Redwood Junction
1740 W. Research Way
Decker Lake
3070 S. Decker Lake Blvd
West Valley Central
2750 W. Lehman Ave
Millcreek
210 W. 3300 South
Meadowbrook
188 W. 3900 South
Murray North
71 W. Fireclay Ave
Murray Central
140 W. Vine St
UTAH TRANSIT AUTHORITY
TRAX
LINES
Blue (701)
(Downtown – Sandy)
Red (703)
(University of Utah – Daybreak)
Green (704)
(Downtown – West Valley City)
Station
Station with park-and-ride
AUGUST 2011
Bingham Junction
7387 S. Bingham Jct Blvd
Fashion Place West
222 W. Winchester St
Historic Gardner
1127 W. 7800 South
Midvale Fort Union
180 W. 7250 South
West Jordan City Center
8021 S. Redwood Rd
Midvale Center
95 W. Center St
2700 West
8351 S. 2700 West
Jordan Valley
3400 W. 8600 South
Historic Sandy
165 E. 9000 South
4800 West
4773 W. Old Bingham Hwy
Sandy Expo
115 E. 9400 South
5600 West
5651 W. Old Bingham Hwy
Sandy Civic Center
115 E. Sego Lily Dr
South Jordan Parkway
10605 S. Grandville Ave
Daybreak Parkway
11405 S. Grandville Ave

L5: Visual Encoding

REQUIRED READING

3

Visual Encoding Principles

The previous chapter presented a taxonomy of data types; now we discuss how to encode these types with a visual representation. This chapter begins with discussion of human perception as a system for making relative rather than absolute judgements. It continues with the image theory of marks and visual channels, and a discussion of channel types. The chapter presents a ranking of channels according to the data type that they are used to encode, and discusses the concepts of expressiveness and expressiveness. It continues with ways to measure channel effectiveness in terms of accuracy, discriminability, and separability, and ability to provide visual popout. The chapter then discusses the characteristics of each channel, including planar spatial position, color, size, tilt and angle, shape, and stipple. It ends with a discussion of the difficulties of 3D depth coding.

3.1 Relative vs. Absolute Judgements

The human perceptual system is fundamentally based on relative judgements, not absolute ones. **Weber's Law** states that the amount of difference that we can detect is relative to the context between the two things, not an absolute quantity. [1] For instance, the amount of length difference we can detect is a percentage of the object's length.

This principle holds true for all sensory modalities. The fact that our

Open PDF in Preview

Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design

Jeffrey Heer and Michael Bostock
Computer Science Department
Stanford University
{jheer, mbostock}@cs.stanford.edu

ABSTRACT
Understanding perception is critical to effective visualization design. With its low cost and scalability, crowdsourcing presents an attractive option for evaluating the large design space of visualizations; however, it first requires validation. In this paper, we assess the viability of Amazon's Mechanical Turk as a platform for graphical perception experiments. We replicate previous studies of spatial encoding and luminance contrast and compare our results. We also conduct new experiments on rectangular area perception (as in treemaps or cartograms) and on chart size and gridline spacing. Our results demonstrate that crowdsourced perception experiments are viable and contribute new insights for visualization design. Lastly, we report cost and performance data from our experiments and distill recommendations for the design of crowdsourced studies.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces—Evaluation/Methodology

General Terms: Experimentation, Human Factors.

Keywords: Information visualization, graphical perception, user study, evaluation, Mechanical Turk, crowdsourcing

for ecological validity. Crowdsourced experiments may also substantially reduce both the cost and time to result.

Unfortunately, crowdsourcing introduces new concerns to be addressed before it is credible. Some concerns, such as ecological validity, subject motivation and expertise, apply to any study and have been previously investigated [13, 14, 23]; others, such as display configuration and viewing environment, are specific to visual perception. Crowdsourced perception experiments lack control over many experimental conditions, including display type and size, lighting, and subjects' viewing distance and angle. This loss of control inevitably limits the scope of experiments that reliably can be run. However, there likely remains a substantial subclass of perception experiments for which crowdsourcing can provide reliable empirical data to inform visualization design.

In this work, we investigate if crowdsourced experiments insensitive to environmental context are an adequate tool for graphical perception research. We assess the feasibility of using Amazon's Mechanical Turk to evaluate visualizations and then use these methods to gain new insights into visualization design. We make three primary contributions: